Medical Informatics Europe 2002
G Surján et al. (eds)
IOS Press 2002

Tools for Constructive Assessment of bids to a Call for Tender - some experiences

Jytte BRENDER[1], Peter McNAIR[2]

1) Aalborg University, Institute of Health Science and Technology, and V-CHI, DK-9220 Aalborg E, Denmark; jytte.brender@v-chi.dk.
2) Hvidovre Hospital, Centre of Laboratories, DK-2650 Hvidovre, Denmark.

Abstract. The paper describes the experiences from a full-scale case study applying a number of novel assessment techniques for selecting among income bids at a call for tender, based on a User Requirements Document that comprises non-prescriptive, goal-oriented requirements.

1. Introduction

A cycle is being closed: The starting point was a study of a 'tailor-made' Laboratory Information System (LIS) in daily operation, see [1,2], exploring why the users did not get what they originally aimed at. The conclusion was that a lot of the problems could be explained by problems in the approach and role of the User Requirements Specification Document (URD). The consequential recommendations were recently applied at the formulation of an URD aimed at a call for tender for the purchase of a LIS [2]. The essentials of the recommendations are: 1) to prepare the URD within the users' terminology rather than on technology's grounds; 2) to extend the URD (and consequently the contract) with a description of strategic aims and procedural ideas and subsequently extend the warranty to include a functionality assessment rather than only a technical verification of functions.

One implication of the recommendations is that the functional requirements are formulated in a goal-oriented way rather than as specific prescriptive details regarding the actual design of the solution. This was agreed upon by the users based on the perspective that a commercial solution was aimed at, and that there might be multiple, equally valid solutions. To support this approach, the bidders were asked to describe their strategic considerations, principles and ideas or mechanisms behind particular solutions, such as the principles behind the design of the user interface, etc., see [2].

The main tools found in the literature to assess bids constitute different kinds of rating and weighting tools, like in the paradigmatic recommendations of for instance [3, 4] or a Balanced Scorecard approach like [5]. These approaches, however, were judged inadequate for our purpose, among others a) because the requirements are not independent, b) because (as also experienced by [6]) none of the bids provided fully completed solutions that at all could be rated and weighted, and c) because they don't provide a holistic management overview of each solution.

This study, which is implemented as case-based action research (see [7]), describes experiences and lessons from using a number of dedicated tools at the constructive

assessment process, leading to the selection of one winning bid. A parallel study goes into detail with this process; see [7].

2. The Case

The organisation within which the current case study takes place is the Copenhagen Hospital Corporation (H:S), comprising 6 hospitals, see further details in [7]. The URD that served as the basis for the call for tender was structured into three levels and included 387 requirement items, in agreement with the ideas, philosophy and strategy inherent in an ISO 900x approach, see [2]: a) a strategic level concerned with the circumscription of the task (leading to the objectives), the conditions, and policy once established and strategy for the solution, including selection of the overall approach and method to be applied; b) a tactical level, concerned with operationalisation of the methods and approaches on the concrete case, including guidelines for the actual solution; and c) an operational level that implements the two previous levels in terms of specific prescriptions, tools and procedures.

The purchase process was implemented as an international call for tender, at which six bids were received and found valid. At the point in time of writing, the purchase has started the implementation of the selected solution by entering a contractual relation with one of the bidding vendors.

3. Methods: assessment tools for the selection process

The perspective on assessment is that an assessment process is non-deterministic and that the users' responsibility for the operation of the future IT-based system must be emphasised and enforced by the assessment approach. This, the definition of the concept of users and the implied strategy are described in detail in [7]. Moreover, the strategy was to apply a constructive assessment methodology, supporting the decision-making process in an objective, incremental fashion and with information-needs driven development of assessment methods and techniques to support the decision-making in a way that interfere minimally with their decision-making process. The tactical approach was:

a) to prepare metrics and implement them as spreadsheet-based tools that would synthesise a large volume of data into fairly dense information;
b) to prepare the global decision-making basis in a hierarchical way that would enable a stepwise progression into increasingly detailed documentation, in case the decision-makers so wished;
c) to prepare and refine assessment techniques in an objective fashion.

Initially, an assessment model was prepared, which included a complete set of assessment metrics, measures and tools, exploiting the options inherent within the features of the Traceability Matrix (TM). The TM comprises an Excel spreadsheet table with the 387 hierarchically numbered requirements of the URD. Each requirement had to be filled for each bid with a score from 0 ('cannot and will not be satisfied') to 3 ('fulfilled by bidder's proposed solution'), where 2 means 'adaptations needed' (additional price requested) and 1 means 'new development is required' (additional price requested). The assumptions for application of the scores were not fulfilled in all cases, as shall be seen.

A number of metrics was calculated on the basis of the six bids: a) each bid on its own (vertical metrics), and b) averaging of the scores for each of the individual requirement items across all bids received (horizontal metrics). The horizontal metrics summarise the information on requirements fulfilment at different levels of importance for the user. A number of horizontal metrics were applied to judge the fairness of each requirement within

the URD. The vertical metrics were used to characterise the individual bids, in terms of different aspects of seriousness, like reservations, misunderstandings, sales talk, etc.

The horizontal score on 'Reasonability', see Table I, indicated the fraction of bids that had a compliance score of either 3 or 2. This metric was used to reveal whether any requirement was unreasonable. Similarly, the metric 'Max degree of fulfilment' indicates something about the feasibility of a requirement.

Table I: Example for Requirement 1.4.1.1. of the horizontal metrics applied on the Traceability Matrix.

REQUIREMENT			Horizontal metrics			
Number	Item	Type	Reason-ability of requirement	Max degree of fulfilment	Median add-on price	Max / mean add-on price
1.4.3.1.	(R) The LIS must conform to national legislation and regulations related to the mission of the LABs and to the LIS' role within the LABs.	R	83%	3	0	0

The methods and metrics initially developed reflect a global approach of a two-step selection process to progress from analysis of the (degree of) requirements fulfilment to a preliminary analysis of the (degree of) objectives fulfilment, combined with a risk and consequence analysis, concluding with a phase of clarification and contract formulation. This two-step selection process later was changed into a principle of exclusion [7].

Table II: An example of two of the vertical metrics for one bid. The 'OK' sign is a control-check on data correctness. 'N/A' means 'not answered' by the bidder. ('R' and 'D' constitutes respectively mandatory and desirable requirements)

Bid C OK **Subjective aspects**

Summary of Bid C's data **Reservations / Lack of commitment**

Type R/D	Degree of fulfilment, % distribution					Total add-on price
	0	1	2	3	N/A	
R	16%	0%	1%	83%	0%	0
D	19%	0%	0%	81%	0%	0
Total	16%	0%	1%	83%	0%	0

Type R/D	% distribution					Total
	0	1	2	3	N/A	
% of R	29%	0%	40%	4%	20%	7%
% of D	0%	100%	0%	2%	0%	3%
Total	1%	0%	1%	3%	1%	6%

4. Some results and lessons achieved

An experience was that the bidders did not fill in the requested information on add-on price for the required, non-fulfilled development of features, which impaired some of the metrics; see for instance the rightmost two columns in Table I and the column under the 'OK' mark in Table II. These metrics were thus excluded from the case study.

The score on 'Reasonability' in Table I ranged from 17% (only one requirement) to 100% of the bids. The 'Max Degree of Fulfilment' was 3 for all requirements. Thus, it is concluded that none of the requirements are absolutely out of reach. The requirement with a reasonability score of 17% was concerned with the access to the source code. More interesting is that an obvious requirement like the fulfilment of the national legislation was abstained in one bid.

Vertical metrics are applied to synthesise qualities and characteristics of the individual bids; see Table II. This was done not only for the distribution on the compliance score, but also to overview the total number of, for instance, misunderstandings, lack of commitment, 'sales' talk, to-be-clarified and specific comments. Each vendors' own statement of

fulfilment was compared with the consultants' initial interpretation of the same; see the example in Table III.

The distribution of the different deviations from the desired compliance score (a measure of 3) was fairly evenly distributed on the Strategic level, the Tactical Level and the Operational Level requirements. There was no singular and obvious candidate bid, but several with very high scores in the interpreted requirements fulfilment. The interpreted, global requirements fulfilment ranged from 60% to 92% with four, comparable bids ranging between 84% and 92% fulfilment.

Table III: Contingency Table showing differences between the Bidder's own statement and H:S's judgment of fulfilment.

A	Bidder	Bidder's statement					Total
		0	1	2	3	n/a	
H:S' judgement	0	0%	1%	2%	0%	0%	2%
	1	0%	4%	10%	3%	0%	18%
	2	0%	0%	5%	1%	0%	6%
	3	0%	1%	8%	59%	0%	68%
	n/a	0%	1%	2%	4%	0%	7%
	Total	0%	7%	26%	67%	0%	100%

A presentation round with demonstration of the solutions was accomplished, based on four common fairly detailed scenarios that addressed tricky aspects of the laboratory practice and the federation of parts of the analytical production, and a number of bid-specific elaborating questions. This presentation round revealed that none of the 6 solutions offered was fully implemented as offered, nor would they be able to fulfil all requirements.

The early stage focussing on 3 out of the 6 bids received was based entirely on the TM with executive summaries (in Tables I-III) and the presentation rounds.

As a supplement of the overview tables I and II, the levels of agreement were calculated for the Contingency Table as overall scores on the hatched fields (in the shown case 69% of full agreement (darkest hatching), and 88%, where only small disagreements occur (all hatched fields). The highest scores of the two were 93% respectively 95% for one bid, indicating a really good impression of an honest bid, not covering up or trying to give an impression of a better degree of requirements fulfilment. The worst of the bids had agreements at the level of 44% and 50%, i.e. a lot of diverging opinions between the bidder and the customer representatives, indicating a less trustworthy bid. The rest of the bids were in the range of 84% and 92% for both agreement indicators, indicating bids with a fairly even level of quality.

Table IV: Example of the Colour Palette. Rows corresponds to individual (or a synthesis of) requirement items. An almost black background indicates an item not yet clarified. Dark grey indicates an insufficient clarification, and colours indicate different degrees of unsatisfactory solutions (red = 'highly significant insufficiency', yellow = 'serious lacking, probably with an option for a solution', and green = 'shortcoming that may be overcome'). Blue text indicates an extraordinary good solution and green text a good solution.

Functional aspect	URD Requirement Item	Bid A	Bid B	Bid C	Bid D	Bid E	Bid F
Future viability	conveyor belts	SampleID is a must	existing option			existing option	
	2.1.8. Point-of-care technology	a little cumbersom	via HL7	OK (depend on a concentrator)	via concentrator	to be clarified. Licenses? functionality OK	one licens per POCT
....							

Subsequent decision-making material was synthesised into an executive summary in terms of an extensive table of pro's and con's, coloured according to severity, and therefore given the nickname, 'Colour Palette', see Table IV. The users filled the Colour Palette, while the consultants filled Tables I to III.

More extensive presentations and a hands-on demonstration were accomplished for three bids. The accomplished site visits put emphasis on preparing a problem analysis by means of the technique described in [8], yet simplified. This was used to elaborate the Colour Palette. An experience was that while Tables I to III were invaluable at the initial exclusion of some of the bids, Table IV turned out to be very valuable in all later discussions. It served throughout as a main part of the decision-making basis together with an update of the economic data and plain supportive information of a descriptive nature. Table IV was fairly easy to maintain and keep updated, while Table I to III was cumbersome to keep updated for so many bids and ongoing clarifications, where the actual changes in terms of implications were not sufficiently visible. Consequently, Table I to III were not used at later stages of the study, except for a concluding summary at the final selection of one bid.

A synthesis of all information was included at the final decision, supported with a risk analysis and an analysis of aspects of the future viability of the LIS solutions.

5. Discussion

This paper contributes with the experiences from application of a number of novel assessment techniques at a (first and almost) full-scale case study, applying a non-prescriptive formulation of the requirements in an URD different from most URDs.

The approach and ideas within the assessment methodology have gradually matured in a chain of projects [2], while the present assessment tools are new. A risk of a 'circular inference' bias arises when one develops a method, framework or technique dedicated for a specific (population of) case(s) and applies it on the very same case(s) for verification, as *is* the case in the present study. Consequently, one should hesitate to conclude that the methods and tools applied are more than applicable, although they indeed worked very well for the case at hand. A study of the implementation and daily operation of the LIS within the case and at minimum one more case is needed to be able to indicate the probability of internal and external validity of the methods and tools.

Acknowledgement

The feasibility of this work was enabled through the role as assigned consultant, while Virtual Centre for Health Informatics at Aalborg University funded the connected research activities. The entire LIS-e organisation is thanked for their enthusiastic engagement, creative spirit and for their constructive collaboration throughout, and especially we thank Jens Schou-Christensen, Infonova, for valuable feedback on the methodology and metrics.

References

[1] Brender J, McNair P. User Requirements in a System Development & Evaluation Context. In: Hasman A, Blobel B, Dudeck J, Gell G, Prokosch H-U, eds. Medical Informatics Europe '00. Amsterdam:IOS Press. Studies in Health Technology and Informatics 2000;**77**:203-207.

[2] Brender J, McNair P. User requirements specifications: a hierarchical structure covering strategical, tactical and operational requirements. Int. J. of Med. Informatics 2001;**64(2-3)**:83-98.

[3] Jeanrenaud A, Romanazzi P. Software product evaluation metrics: a methodological approach. In: Ross M, Brebbia CA, Staples G, Stapleton J, eds. Proceedings of the Software Quality Management II Conference "Building Quality into Software". Southampton:Comp. Mech. Publications. 1994, p.59-69.

[4] Celli M, Ryberg DE, Leaderman AV. Supporting CPR development with the Commercial Off-the-Shelf Systems Evaluation Technique: Defining requirements, setting priorities, and evaluating choices. J Healthcare Inform Manage 1998;12(4):11-19.
[5] Niss K. The use of the Balanced ScoreCard (BSC) in the model for investment and evaluation of medical information systems. In: Kokol P, Zupan B, Stare J, Premik M, Engelbrecht R, eds. Medical Informatics Europe '99. Amsterdam:IOS Press. Studies in Health Technology and Informatics 1999;**68**:110-114.
[6] Madjaric M, Leodolter W, Gell G. HIS purchase project: Preliminary report. In: Kokol P, Zupan B, Stare J, Premik M, Engelbrecht R, eds. Medical Informatics Europe '99. Amsterdam:IOS Press. Studies in Health Technology and Informatics 1999;**68**:115-120.
[7] Brender J, McNair. A case study on Constructive Assessment of bids to a Call for Tender. (this proceedings)
[8] Handbook for Objectives-Oriented Planning. 2nd edition. Oslo:Norwegian Agency for Development Cooperation. 1992.